



Australian Bureau of Statistics

1352.0.55.098 - Research Paper: Generalised Linear Models with Probabilistically Linked Data (Methodology Advisory Committee), November 2008

Latest ISSUE Released at 11:30 AM (CANBERRA TIME) 26/02/2009

Summary

About this Release

The Australian Bureau of Statistics has embarked on the Census Data Enhancement project, the key feature of which is to create a Statistical Longitudinal Census Dataset (SLCD) based on a random sample of 5% of person records from the 2006 Census. These will be linked to person records from 2011 and subsequent Censuses without using names and addresses as linking variables. The SLCD will provide a substantial opportunity for longitudinal analysis to see how people and their families change with time, while maintaining the ABS' strong commitment to the confidentiality of its Census respondents. Since a unique person identifier will not be available, some links will be incorrect, so some linked Census records will not correspond to the same individual. The ABS has conducted a quality study to assess the feasibility of forming the SLCD in this way and its likely quality. Part of the assessment has been to fit generalised linear models to longitudinal linked data. This paper describes and implements a method of adjusting regression coefficients in such models to account for incorrect links. Empirical results show that the adjustment method works well, especially as the number of incorrect links increases. Empirical findings also suggest that a possibly more significant source of error arises when certain sub-populations are underrepresented in the linked data set.

© Commonwealth of Australia

All data and other material produced by the Australian Bureau of Statistics (ABS) constitutes Commonwealth copyright administered by the ABS. The ABS reserves the right to set out the terms and conditions for the use of such material. Unless otherwise noted, all material on this website – except the ABS logo, the Commonwealth Coat of Arms, and any material protected by a trade mark – is licensed under a Creative Commons Attribution 2.5 Australia licence